STATISTICS AND PROBABILITY
First Semester 2025-2026

Mid-Semester Exam
Wednesday, October 8, 2025
A 120-minute open book exam

Problem 1 (SO-PI a1, a2, a3; 20%)

A structural engineer is comparing the compressive strength of concrete from two different suppliers (Supplier A and Supplier B) to determine which one to use for a bridge construction project. Fifteen concrete cylinder samples were tested from each supplier after 28 days of curing. The compressive strength data (in MPa) are as follows.

Supplier A: 35, 33, 36, 31, 37, 35, 33, 34, 36, 35, 37, 34, 36, 35, 34
Supplier B: 28, 31, 29, 35, 30, 33, 27, 34, 31, 29, 30, 32, 28, 31, 33

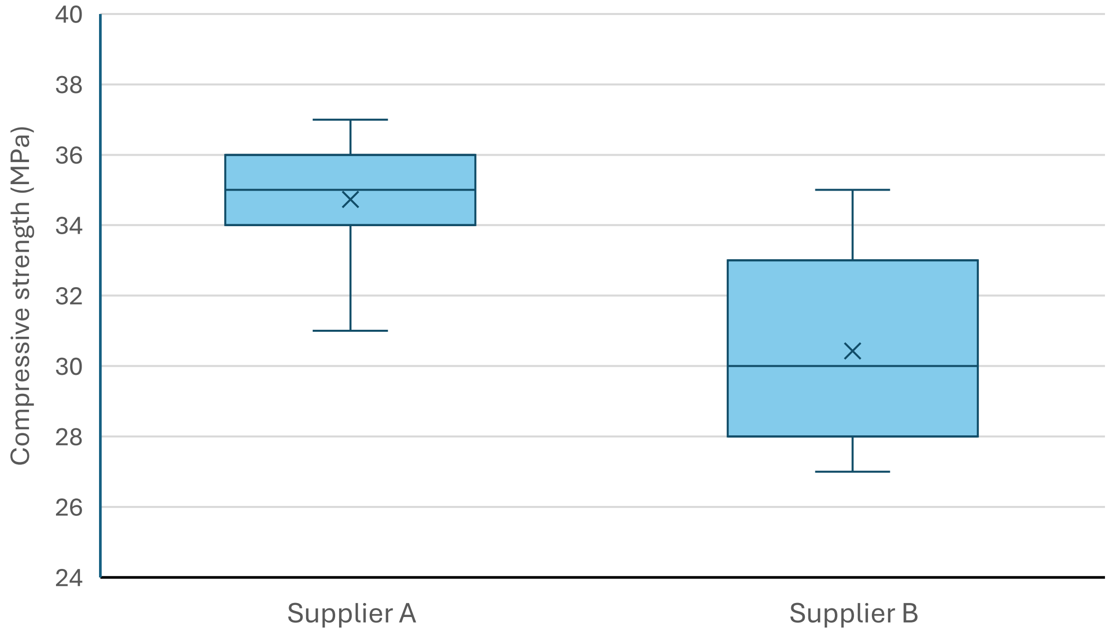
a. For each supplier, calculate the following descriptive statistics: mean, median, first quartile, third quartile, interquartile range, minimum, and maximum values.
b. Construct side-by-side boxplots for both suppliers on the same graph. Compare the two distributions of compressive strength from different suppliers and give your comments.

Answer

a. You can get the descriptive statistics of the data by using your calculator without sorting the data in ascending order. The following table shows these statistics. All figures are in megapascals (MPa).

	mean	median	first quartile	third quartile	inter-quartile range	mini-mum	maxi-mum
Supplier A	34.7	35	34	36	2	31	37
Supplier B	30.7	31	31	35	4	27	35

b. The boxplots in the following page show that Supplier A delivers higher mean and median values of compressive strength, shorter range of interquartile range, and wider minimum-maximum range than Supplier B does. Higher means and median values advise higher quality of concrete. Shorter interquartile range suggests more spread in the concrete quality. The wide range of minimum-maximum range reiterates this spread.



Problem 2 (SO-PI a1, a2, a3; 15%)

A bridge construction project crosses a river. After several years, segments were inspected and it is found that:

40% show structural damage,
30% show environmental damage,
15% show both types of damage.

a. Suppose that a segment was inspected, what is the probability that a segment has at least one type of damage?
b. If a segment shows damage, what is the probability that it is environmental damage?
c. What is the probability that a segment shows neither damage?

Answer

Suppose that event A expresses structural damage, event B states environmental damage. One has

$$\text{prob}(A) = 0.40, \text{prob}(B) = 0.30, \text{and prob}(A \cap B) = 0.15$$

The following Venn diagram gives clear representation of the situation. Looking at this diagram, the problems are easily solved.

a. The probability that a segment has at least one type of damage is 55%.

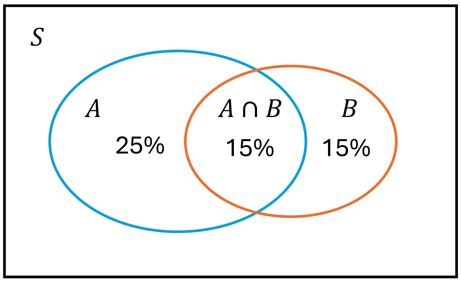
$$\text{prob(at least a damage)} = \text{prob}(A \cup B) = \text{prob}(A) + \text{prob}(B) - \text{prob}(A \cap B) = 55\%$$

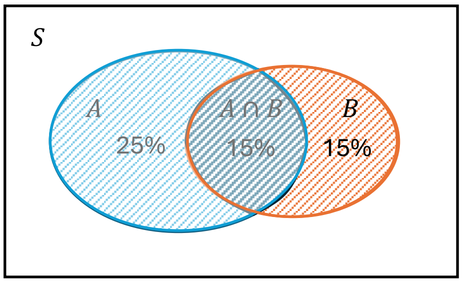
b. The probability of environmental damage when damage occurs is 54.54%.

$$\text{prob(environmental damage|damage has occurred)} = \frac{\text{prob}(B)}{\text{prob}(A \cup B)} = \frac{30\%}{55\%} = 54.54\%$$

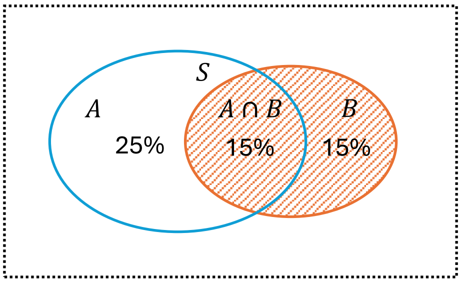
c. The probability that a segment shows neither damage is 45%.

$$\text{prob(no damage)} = 1 - \text{prob}(A \cup B) = (100 - 55)\% = 45\%$$

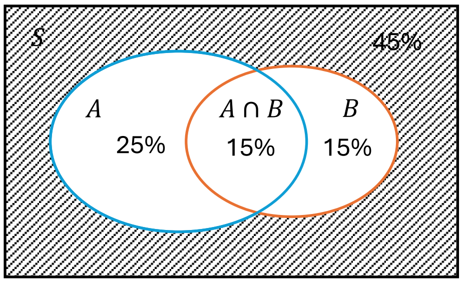




(a)



(b)



(c)

Problem 3 (SO-PI a1, a2, a3; 15%)

The number of trucks arriving at a construction site per hour follows a discrete probability distribution.

The number of trucks (X)	0	1	2	3	4
$P(X)$	0.10	0.25	0.35	0.20	0.10

a. Verify that the probability distribution above is valid.
b. What is the probability that at least 2 trucks will arrive in an hour?
c. Calculate the expected number of trucks per hour.

Answer

a. The probability distribution is valid since the total probability of trucks arriving at a construction site per hour is unity.

$$\text{prob(all possible occurrence)} = 0.10 + 0.25 + 0.35 + 0.20 + 0.10 = 1$$

b. The probability that at least 2 trucks will arrive in an hour is 0.70 (70%).

$$\text{prob(at least 2 trucks in an hour)} = 0.10 + 0.25 + 0.35 = 0.70$$

c. The expected number of trucks per hour is 2.

$$E(\text{number of trucks per hour}) = 0.10 \times 0 + 0.25 \times 1 + 0.35 \times 2 + 0.20 \times 3 + 0.10 \times 4 = 2$$

Problem 4 (SO-PI a1, a2, a3; 25%)

Traffic engineers counted the number of vehicles passing through an intersection. On average, 3.5 vehicles pass through per minute during off-peak hours.

a. What is the probability that exactly 4 vehicles will pass in a given minute?
b. What is the probability that no vehicle passes in a given minute?
c. What is the probability that more than 5 vehicles pass in a given minute?
d. Calculate the probability that at least 2 vehicles pass in a 30-second interval.

Answer

This is a Poisson distribution, which is a discrete process on a continuous time scale. Let $\lambda = 3.5$ vehicles per minute.

a. The probability that exactly 4 vehicles will pass in a given minute is 19%.

$$f_X(x;\lambda) = \frac{\lambda^x e^{-\lambda}}{x!} \Leftrightarrow f_X(4; 3.5) = \frac{3.5^4 e^{-3.5}}{4!} = 0.1888 = 18.88\% \approx 19\%$$

b. The probability that no vehicle passes in a given minute is 3%.

$$f_X(x;\lambda) = \frac{\lambda^x e^{-\lambda}}{x!} \Leftrightarrow f_X(0; 3.5) = \frac{3.5^0 e^{-3.5}}{0!} = 0.03 = 3\%$$

c. The probability that more than 5 vehicles pass in a given minute is 27.5%

$$1 - F_X(x;\lambda) = 1 - \sum_{i=0}^{x} \frac{\lambda^i e^{-\lambda}}{i!}$$

$$1 - F_X(4; 3.5) = 1 - \sum_{i=0}^{4} \frac{3.5^i e^{-3.5}}{i!} = 1 - (0.0302 + 0.1057 + 0.1850 + 0.2158 + 0.1888)$$
$$= 1 - 0.7254 = 0.2746 \approx 27.5\%$$

d. Since the time interval changes, so does the parameter λ. In a 30-second interval, one has $\lambda = 3.5 \times 30/60 = 1.75$ vehicles pass through the intersection. Therefore, the probability that at least 2 vehicles pass through the intersection in 30 seconds is 52.2%.

$$1 - F_X(1; 1.75) = 1 - \sum_{i=0}^{1} \frac{1.75^i e^{-1.75}}{i!} = 1 - (0.1738 + 0.3041) = 1 - 0.4779 = 0.5221$$
$$\approx 52.2\%$$

Problem 5 (SO-PI a1, a2, a3; 25%)

The concentration of lead in soil samples from a contaminated industrial site follows a normal distribution with a mean of 450 ppm (parts per million) and a standard deviation of 80 ppm.

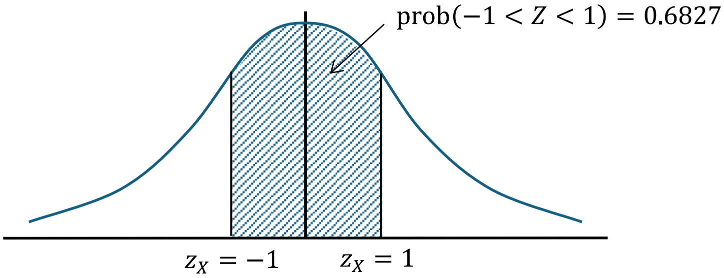
a. What percentage of samples has lead concentration between 370 ppm and 530 ppm?
b. What is the probability that a randomly selected sample has lead concentrations exceeding 600 ppm?

c. Find the concentration that separates the most contaminated 5% of samples from the rest.
d. If the remediation threshold is 300 ppm, what proportion of samples require cleanup?

Answer

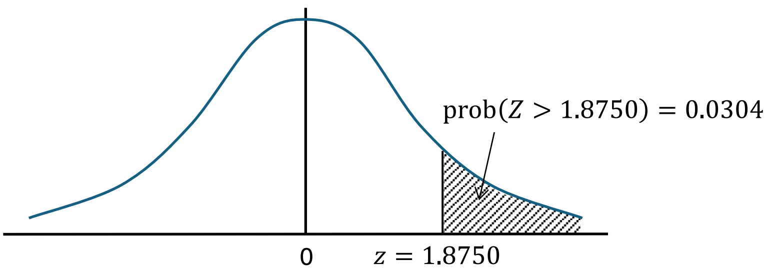
Let X is the concentration of lead in ppm unit. One has a normal distribution whose parameters are $\bar{X} = 450$ and $s_X = 80$.

a. The percentage of samples that has lead concentrations between 370 ppm and 530 ppm is 68.27%.



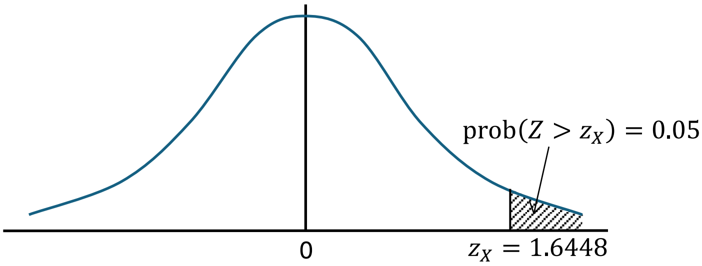
$$
\begin{aligned}
\text{prob}(370 < X < 530) &= \text{prob}\left(\frac{370-450}{80} < Z_X < \frac{530-450}{80}\right) \\
&= \text{prob}\left(\frac{370-450}{80} < Z_X < \frac{530-450}{80}\right) = \text{prob}(-1 < Z_X < 1) \\
&= \text{prob}(Z_X < 1) - \text{prob}(Z_X < -1) = 0.8413 - 0.1587 = 0.6827 = 68.27\%
\end{aligned}
$$

b. The probability that a randomly selected sample has lead concentrations exceeding 600 ppm is 3%.



$$
\begin{aligned}
\text{prob}(X > 600) = 1 - \text{prob}(X < 600) &= 1 - \text{prob}\left(Z_X < \frac{600-450}{80}\right) \\
&= 1 - \text{prob}(Z_X < 1.8750) = 1 - 0.9696 = 0.0304 = 3\%
\end{aligned}
$$

c. The concentration that separates the most contaminated 5% of samples from the rest is 582 ppm.

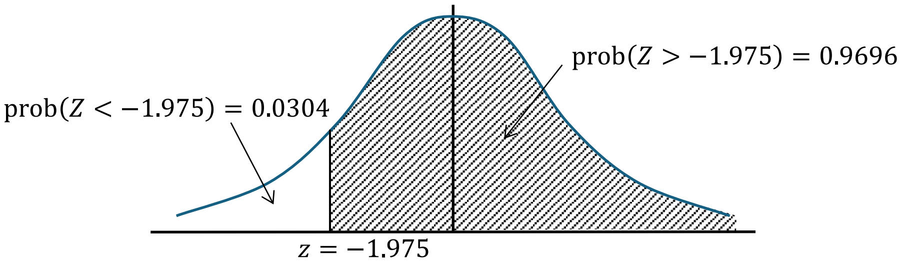


$$\text{prob}(X > x) = 5\% \Leftrightarrow \text{prob}(X < x) = 100\% - 5\% = 95\%$$

$$\text{prob}(Z < z_X) = 95\% \Leftrightarrow z_X = 1.6448$$

$$x = \bar{X} + Z_X s_X = 450 + 1.6448 \times 80 = 582$$

d. The proportion of samples that require cleanup is 97% if the remediation threshold is 300 ppm.



$$\text{prob}(X > 300) = 1 - \text{prob}(X < 300) = 1 - \text{prob}\left(Z_X < \frac{300 - 450}{80}\right)$$
$$= 1 - \text{prob}(Z_X < -1.975) = 1 - 0.0304 = 0.9696 = 97\%$$

-o0o-